

- A speech recognition device, comprising:
 an I/O device for accepting a voice stream;
- a frequency domain converter communicating with said I/O device, said frequency domain converter converting said voice stream from a time domain to a frequency domain and generating a plurality of frequency domain outputs;
- a frequency domain output storage communicating with said frequency domain converter, said frequency domain output storage comprising at least two frequency spectrum frame storages for storing at least a current frequency spectrum frame and a previous frequency spectrum frame, with a frequency spectrum frame storage of said at least two frequency spectrum frame storages comprising a plurality of frequency bins storing said plurality of frequency domain outputs;
- a processor communicating with said plurality of frequency bins;
 - a memory communicating with said processor;
- a frequency spectrum difference storage in said memory, with said frequency spectrum difference storage storing one or more frequency spectrum differences calculated as a difference between said current frequency spectrum frame and said previous frequency spectrum frame;

at least one feature storage in said memory for storing at least one feature extracted from said voice stream;

at least one transneme table in said memory, with said at least one transneme table including a plurality of transneme table entries and with a transneme table entry of said plurality of transneme table entries mapping a predetermined frequency spectrum difference to at least one predetermined transneme of a predetermined verbal language;

at least one mappings storage in said memory, with said at least one mappings storage storing one or more found transnemes;

at least one transneme to-vocabulary database in said memory, with said at least one transneme-to-vocabulary database mapping a set of one or more found transnemes to at least one speech unit of said predetermined verbal language; and

at least one voice stream representation storage in said memory, with said at least one voice stream representation storage storing a voice stream representation created from said one or more found transpemes;

wherein said speech recognition device calculates a frequency spectrum difference between a current frequency spectrum frame and a previous frequency spectrum frame, maps said frequency spectrum difference to a transneme table, and converts said frequency spectrum difference to a transneme if said frequency spectrum difference is greater than a predetermined difference threshold, and creates a digital voice stream representation of said voice stream from one or more transnemes thus produced.

- 2. The speech recognition device of claim 1, wherein said voice stream is accepted as a digital voice stream.
- 3. The speech recognition device of claim 1, wherein said voice stream is compressed.
- 4. The speech recognition device of claim 1, wherein said I/O device comprises a microphone.
- 5. The speech recognition device of claim 1, wherein said I/O device comprises a wireless receiver.
- 6. The speech recognition device of claim 1, wherein said I/O device comprises a digital network interface.
- 7. The speech recognition device of claim 1, wherein said I/O device comprises an analog network interface.
- 8. The speech recognition device of claim 1, wherein said frequency domain converter is a Fourier transform device.
- 9. The speech recognition device of claim 1, wherein said frequency domain converter is a filter bank comprising a plurality of predetermined filters.

- 10. The speech recognition device of claim 1, wherein said frequency domain output storage is in said memory.
- 11. The speech recognition device of claim 1, wherein said memory further comprises a feature storage and said processor communicates with said frequency domain output storage and extracts at least one feature from said voice stream in a frequency domain and stores said at least one feature in said feature storage.
- 12. The speech recognition device of claim 1, wherein said memory further comprises a feature storage and said processor communicates with said I/O device and extracts at least one feature from said voice stream in a time domain and stores said at least one feature in said feature storage.
- 13. The speech recognition device of claim 1, wherein said frequency domain converter, said frequency domain output storage, said processor, and said memory are included on a digital signal processing (DSP) chip.
- 14. The speech recognition device of claim 1, wherein said digital voice stream representation comprises a series of symbols.

- 15. The speech recognition device of claim 1, wherein said digital voice stream representation comprises a series of text symbols.
- 16. The speech recognition device of claim 1, wherein said speech recognition device converts and compresses said voice stream into a compressed digital voice stream representation comprising a series of symbols.
- 17. The speech recognition device of claim 1, wherein said speech recognition device converts and compresses said voice stream into a compressed digital voice stream representation and transmits said compressed digital voice stream representation as a series of symbols.
- 18. The speech recognition device of claim 1, wherein said speech recognition device converts and compresses said voice stream into a compressed digital voice stream representation and stores said compressed digital voice stream representation as a series of symbols.
- 19. A method for performing speech recognition on a voice stream, comprising the steps of:

determining one or more candidate transnemes in said voice stream;

mapping said one or more candidate transnemes to a transneme table to convert said one or more candidate transnemes to one or more found transnemes; and

mapping said one or more found transnemes to a transneme-to-vocabulary database to convert said one or more found transnemes to one or more speech units.

- 20. The method of claim 19, wherein said one or more speech units are combined to create a digital voice stream representation of said voice stream.
- 21. The method of claim 19, wherein said one or more speech units are combined to create a digital voice stream representation of said voice stream, with said digital voice stream representation comprising a series of symbols.
- 22. The method of claim 19, wherein said one or more speech units are combined to create a digital voice stream representation of said voice stream, with said digital voice stream representation comprising a series of text symbols.
- 23. The method of claim 19, with said determining step further comprising comparing at least two frequency spectrum frames in a frequency domain in order to determine said one or more candidate transnemes.

- 24. The method of claim 19, wherein said voice stream is compressed by said method into a compressed digital voice stream representation comprising a series of symbols.
- 25. The method of claim 19, wherein said voice stream is compressed by said method into a compressed digital voice stream representation and wherein said method further comprises a step of transmitting said compressed digital voice stream representation as a series ϕf symbols.
- 26. The method of claim 19, wherein said voice stream is compressed by said method into a compressed digital voice stream representation and wherein said method further comprises a step of storing said compressed digital voice stream representation as a series of symbols.
- 27. The method of claim 19, wherein a voice stream in a first verbal language is converted into a voice stream representation in a second language.
- 28. A method for performing speech recognition on a voice stream, comprising the steps of:

calculating a frequency spectrum difference between a current frequency spectrum frame and a previous frequency spectrum frame, with said current frequency spectrum frame and

said previous frequency spectrum frame being in a frequency domain and being separated by a predetermined time interval; and

mapping said frequency spectrum difference to a transneme table to convert said frequency spectrum difference to at least one transneme if said frequency spectrum difference is greater than a predetermined difference threshold;

wherein a digital voice stream representation of said voice stream is created from one of more transnemes thus produced.

- 29. The method of claim 28, further including the steps of: saving tonality level changes of said voice stream; and using said tonality level changes to add punctuation to said voice stream representation.
- 30. The method of claim 28, wherein at least one feature is extracted from said voice stream in a time domain.
- 31. The method of claim 28, wherein at least one feature is mathematically extracted from said voice stream in a frequency domain.
- 32. The method of claim 28, wherein at least one feature is mathematically extracted from said voice stream in a frequency domain, and wherein said voice stream is a compressed voice stream already in said frequency domain.

33. The method of claim 28, further comprising the steps of:

performing a frequency domain transformation on said voice stream upon a predetermined time interval to create said current frequency spectrum frame;

storing said current frequency spectrum frame in a plurality of frequency bins; and

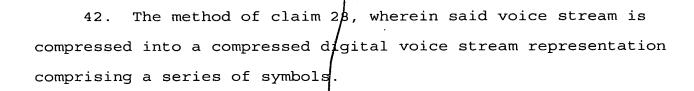
amplitude shifting and frequency shifting said current frequency spectrum frame based on a comparison of a current base frequency of said current frequency spectrum frame to a previous base frequency of a previous frequency spectrum frame.

- 34. The method of claim 28, wherein said predetermined time interval is less than a phoneme in length.
- 35. The method of flaim 28, wherein said predetermined time interval is about ten milliseconds.
- 36. The method of claim 28, wherein said predetermined difference threshold is about 5% of average amplitude of a base frequency bin over a window of less than 100 milliseconds.
- 37. The method of claim 28, further comprising the steps of:

accumulating a predetermined number of transnemes;

performing a lookup of said predetermined number of transnemes against a transneme to-vocabulary database; and matching at least one transneme in said predetermined number of transnemes to at least one speech unit in said transneme-to-vocabulary database.

- 38. The method of claim 37 wherein about ten to about twenty transnemes are accumulated in said predetermined number of transnemes for performing said lookup against said transneme-to-vocabulary database.
- 39. The method of claim 37, with the step of performing a lookup against a transmeme-to-vocabulary database further comprising performing a free-text-search lookup of said predetermined number of transmemes against said transmeme-to-vocabulary database using inverted-index techniques in order to find one or more best-fit mappings of a segment of transmemes in said predetermined number of transmemes to at least one speech unit in said transmeme-to-vocabulary database.
- 40. The method of claim 28, wherein said digital voice stream representation comprises a series of symbols.
- 41. The method of claim 28, wherein said digital voice stream representation comprises a series of text symbols.



- 43. The method of claim 28, wherein said voice stream is compressed by said method into a compressed digital voice stream representation and wherein said method further comprises a step of transmitting said compressed digital voice stream representation as a series of symbols.
- 44. The method of claim 28, wherein said voice stream is compressed by said method into a compressed digital voice stream representation and wherein said method further comprises a step of storing said compressed digital voice stream representation as a series of symbols.
- 45. The method of claim 28, wherein a voice stream in a first verbal language is converted into a voice stream representation in a second language.
- 46. A method for performing speech recognition on a voice stream, comprising the steps of:

performing a frequency domain transformation on said voice stream upon a predetermined time interval to create a current frequency spectrum frame;

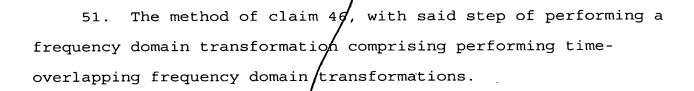
normalizing said current frequency spectrum frame;

calculating a frequency spectrum difference between said current frequency spectrum frame and a previous frequency spectrum frame;

mapping said frequency spectrum difference to a transneme table to convert said frequency spectrum difference to at least one found transneme if said frequency spectrum difference is greater than a predetermined difference threshold; and

creating a digital voice stream representation of said voice stream from one or more found transnemes thus produced.

- 47. The method of claim 46, further including the steps of: saving tonality level changes of said voice stream; and using said tonality level changes to add punctuation to said voice stream representation.
- 48. The method of claim 46, wherein at least one feature is extracted from said voice stream in a time domain.
- 49. The method of claim 46, wherein at least one feature is mathematically extracted from said voice stream in a frequency domain.
- 50. The method of claim 46, wherein at least one feature is mathematically extracted from said voice stream in a frequency domain, and wherein said voice stream is a compressed voice stream already in said frequency domain.



- 52. The method of claim 46, with said step of performing a frequency domain transformation comprising performing a Fourier transformation.
- 53. The method of claim 46, with said step of performing a frequency domain transformation comprising performing time-overlapping frequency domain transformations of a predetermined transformation window about every 5 milliseconds.
- 54. The method of claim 46, with said step of performing a frequency domain transformation comprising performing time-overlapping frequency domain transformations of an about 10 millisecond transformation window about every 5 milliseconds.
- 55. The method of claim 46, further comprising the step of storing said current frequency spectrum frame in a plurality of current frequency bins.
- 56. The method of claim 46, with said step of normalizing comprising normalizing a base frequency of said current frequency

spectrum frame to a base frequency of said previous frequency spectrum frame.

- 57. The method of claim 46, with said step of normalizing comprising frequency shifting said current frequency spectrum frame using an extracted pitch feature.
- 58. The method of claim 46, with said step of normalizing comprising amplitude shifting said current frequency spectrum frame using an extracted volume feature.
- 59. The method of claim 46, with said step of normalizing comprising amplitude shifting and frequency shifting said current frequency spectrum frame based on a comparison of a current base frequency of said current frequency spectrum frame to a previous base frequency of said previous frequency spectrum frame.
- 60. The method of claim 46, further comprising the step of storing said current frequency spectrum frame in a plurality of current frequency bins and with said step of calculating said frequency spectrum difference comprising calculating a plurality of difference values between a plurality of current frequency spectrum frame bin values in said plurality of current frequency bins and a plurality of previous frequency spectrum frame bin values.

- 61. The method of claim 46, wherein said predetermined time interval is less than a phoneme in length.
- 62. The method of claim 46, wherein said predetermined time interval is about ten milliseconds.
- 63. The method of claim 46 wherein said predetermined threshold is about 5% of average amplitude of a base frequency bin over a window of less than 100 milliseconds.
- 64. The method of claim 46, further comprising the steps of:

accumulating a predetermined number of transnemes;

performing a lookup of said predetermined number of

transnemes against a transneme-to-vocabulary database; and

matching at least one transneme in said predetermined number

of transnemes to at least one speech unit in said transneme-to
vocabulary database.

65. The method of claim 64, wherein about ten to about twenty transnemes are accumulated in said predetermined number of transnemes for performing said lookup against said transneme-to-vocabulary database.

- 66. The method of claim 64 with the step of performing a lookup against a transneme-to-vocabulary database further comprising performing a free-text-search lookup of said predetermined number of transnemes against said transneme-to-vocabulary database using inverted-index techniques in order to find one or more best-fit mappings of a segment of transnemes in said predetermined number of transnemes to at least one speech unit in said transneme-to-vocabulary database.
- 67. The method of claim 46, wherein said digital voice stream representation comprises a series of symbols.
- 68. The method of clafm 46, wherein said digital voice stream representation comprises a series of text symbols.
- 69. The method of claim 46, wherein said voice stream is compressed into a compressed digital voice stream representation comprising a series of symbols.
- 70. The method of claim 46, wherein said voice stream is compressed by said method into a compressed digital voice stream representation and wherein said method further comprises a step of transmitting said compressed digital voice stream representation as a series of symbols.

- 71. The method of claim 46, wherein said voice stream is compressed by said method into a compressed digital voice stream representation and wherein said method further comprises a step of storing said compressed digital voice stream representation as a series of symbols.
- 72. The method of claim 46, wherein a voice stream in a first verbal language is converted into a voice stream representation in a second language.